

State of the Art on Diffusion Models for Visual Computing

Year: 2023 | Citations: 148 | Authors: Ryan Po, Wang Yifan, Vladislav Golyanik, Kfir Aberman, J. Barron

Abstract

The field of visual computing is rapidly advancing due to the emergence of generative artificial intelligence (AI), which unlocks unprecedented capabilities for the generation, editing, and reconstruction of images, videos, and 3D scenes. In these domains, diffusion models are the generative AI architecture of choice. Within the last year alone, the literature on diffusion-based tools and applications has seen exponential growth and relevant papers are published across the computer graphics, computer vision, and AI communities with new works appearing daily on arXiv. This rapid growth of the field makes it difficult to keep up with all recent developments. The goal of this state-of-the-art report (STAR) is to introduce the basic mathematical concepts of diffusion models, implementation details and design choices of the popular Stable Diffusion model, as well as overview important aspects of these generative AI tools, including personalization, conditioning, inversion, among others. Moreover, we give a comprehensive overview of the rapidly growing literature on diffusion-based generation and editing, categorized by the type of generated medium, including 2D images, videos, 3D objects, locomotion, and 4D scenes. Finally, we discuss available datasets, metrics, open challenges, and social implications. This STAR provides an intuitive starting point to explore this exciting topic for researchers, artists, and practitioners alike.